ACCEPTED VERSION

Rebecca R. Chao, James J. De Voss and Stephen G. Bell
The efficient and selective catalytic oxidation of para-substituted cinnamic acid derivatives by the cytochrome P450 monooxygenase, CYP199A4
RSC Advances, 2016; 6(60):55286-55297

Published at: http://dx.doi.org/10.1039/c6ra11025h

The efficient and selective catalytic oxidation of *para*-substituted cinnamic acid derivatives by the cytochrome P450 monooxygenase, CYP199A4

Rebecca R. Chao[a], James J. De Voss[b] and Stephen G. Bell[a*]

[a] Department of Chemistry, University of Adelaide, SA 5005, Australia. Fax: +61 8 8303 4380; email: stephen.bell@adelaide.edu.au

[b] School of Chemistry and Molecular Biosciences, University of Queensland, Brisbane, Qld, 4072, Australia.

Abstract

The cytochrome P450 enzyme, CYP199A4, demethylated 4-methoxybenzoic acid, but not 4-methoxyphenylacetic acid, with high product formation activity. The oxidative demethylation of 3-(4-methoxyphenyl)propionic acid was 8-fold more active than 4-methoxyphenylacetic acid and 4-methoxycinnamic acid was efficiently oxidised at 180 min^{-1}. Accordingly the oxidation of cinnamic acid derivatives was investigated in order to determine the potential of CYP199A4 to act as a biocatalyst for this important class of biological molecules. 4-Methoxy- and 4-methyl-cinnamic acids bound tightly to CYP199A4 and both were better substrates for CYP199A4 than cinnamic acid itself. The oxidations were 100% selective for attack at the *para*-methoxy group. Certain dimethoxy substituted cinnamic acids were demethylated more efficiently than 4-methoxycinnamic acid and retained the selectivity for the *para*-methoxy substituent. 4-Isopropylcinnamic acid was hydroxylated and desaturated by CYP199A4. Therefore despite the longer length of the cinnamic acids compared to benzoic acids, a suitable *para* substituent was still essential for optimal substrate binding and activity. In general molecules with a *para*-substituted alkyl- and alkyloxy-cinnamic acid structure were a good fit for the active site of the CYP199A4 enzyme and as a consequence are efficiently and selectively oxidised. Whole-cell oxidations resulted in high yields of product and CYP199A4 could be developed for applications in the biocatalytic oxidation of cinnamic acid derivatives and related phenylpropanoids.

Introduction

The heme-dependent cytochrome P450 (CYP) monooxygenase enzymes primarily catalyse the insertion of an oxygen atom derived from dioxygen into the carbon-hydrogen bonds of organic molecules.[1-3] CYP199A4, from *Rhodopseudomonas palustris* strain HaA2, is able to selectively demethylate 4-methoxybenzoic acid.[4, 5] It is also able to hydroxylate and desaturate *para*-substituted alkyl benzoic acids such as 4-ethyl- and 4-isopropyl-benzoic acids to form a mixture of products.[4, 6] These activities are supported by a class I electron transfer chain consisting of a [2Fe-2S] ferredoxin (HaPux) and a flavin-dependent ferredoxin reductase (HaPuR) which mediates heme reduction by NADH. When CYP199A4 was paired with these electron transfer partners, the activities of oxidative demethylation, hydroxylation and desaturation of *para*-substituted benzoic acids are high.[4, 7, 8]

Mechanistically, C–H bond hydroxylation occurs via hydrogen abstraction by Compound I, followed by oxygen rebound to give the hydroxylated organic.[2, 9] *O*-demethylation of 4-methoxybenzoic acid by CYP199A4 occurs via the same pathway, initially yielding a hemiacetal which spontaneously decomposes to yield 4-hydroxybenzoic acid and formaldehyde.[2, 6, 10, 11] The alkene desaturation product is hypothesised to arise from initial hydrogen atom abstraction followed by either a second hydrogen atom abstraction from an adjacent carbon or one-electron oxidation by Compound II. The latter pathway generates a carbocation which then loses a proton to form the alkene.[2, 6]

A detailed knowledge of the types of substrates that can bind in the enzyme active site and the interactions responsible for binding is of paramount importance in achieving the activity and turnover numbers required for synthetic applications.[12-15] The stable, soluble nature and high activity of the CYP199A4 enzyme means it is well suited for larger scale processes involving C–H bond oxidation such as hydroxylation and oxidative demethylation.[4, 6, 16-18] Furthermore the enzyme could be engineered via rational mutagenesis or directed evolution to enhance its activity and

broaden its substrate range.[5, 19]

Altering the benzene ring, the carboxylate group or the removal of the methoxy substituent of 4-methoxybenzoic acid all reduced the substrate binding and monooxygenase activity of CYP199A4.[4, 6, 20, 21] Replacement of the methoxy group with alternate substituents is tolerated as are small additional functional groups on the benzene ring.[6] We have shown that CYP199A4 can bind and *O*-demethylate 3,4-dimethoxybenzoic acid with total regioselectivity at the 4-methoxy group and demethenylate 3,4-methylenedioxybenzoic acid.[4, 20] Others have reported that the CYP199A subfamily member, CYP199A2 from *R. palustris* CGA009, which shares > 85% sequence identity with CYP199A4, can catalyse the oxidation of larger substrates such as 2-naphthoic acid, indole-6-carboxylic acid and 3-hydroxy- and 4-hydroxy-cinnamic acids (*m*- and *p*-coumaric acids, respectively).[4, 16-18, 22-24] Mutant forms of CYP199A2 have been shown to enhance the biocatalytic properties of CYP199A2, allowing the whole-cell oxidation of cinnamic acid, benzyl alcohols and phenols using an *E. coli* system.[19] However, detailed *in vitro* analysis of the substrate binding interactions and enzyme activity with these substrates has not been reported.

Several crystal structures of substrate-free and substrate-bound forms of CYP199A4 and CYP199A2 have been determined.[5, 6, 10] In the substrate-bound forms (PDB codes; 4DO1, 4EGM and 4EGN), the substituent *para* to the carboxylate is oriented towards the heme-iron which is consistent with exclusive attack at this position. The carboxylate group of the substrate interacts with the polar side chains of Arg92, Ser95, Ser244 and, via a bridging water molecule, Arg243 (Fig. 1). In addition, hydrophobic interactions with the benzene ring help to position the substrate. The crystal structures of CYP199A4 in complex with the larger indole-6-carboxylic acid and 2-naphthoic acid substrates have also been solved (PDB codes; 4EGO and 4EGP). In the indole-6-carboxylic acid-bound structure, the benzene rings, the carboxylate groups and the bridging water molecules are at virtually identical locations to those observed with 4-methoxybenzoic acid.[5]

Although 2-naphthoic acid is similar in size to indole-6-carboxylic acid, its crystallographically observed binding mode is significantly different. The carboxylate group is located *ca*. 1.9 Å closer to the N-terminal end of the I helix. The water molecule found in the 4-methoxybenzoic acid-bound form, which interacts with the substrate and Arg243, is displaced and the carboxylate group of 2-naphthoic acid interacts directly with Arg243.[5] The hydrogen bonds between the carboxylate oxygen atom and the side chains of Ser95 and Ser244 are maintained. ITo preserve these interactions, the alcohol group of Ser95 undergoes a significant positional shift. Together, these structures show that CYP199A4 is able to accommodate larger substrates than benzoic acids.[5] By way of contrast, it has been shown that small divergences from 4-methoxybenzoic acid structure can result in a dramatic reduction in the enzyme activity and in the affinity for the substrate. For example, the activity of CYP199A4 with 4-methoxyphenylacetic acid is 240-fold lower than that with 4-methoxybenzoic acid and there is a three orders of magnitude reduction in the substrate binding affinity.[21]

Cinnamic acids belong to the phenylpropanoid family of natural products and have a variety of biological activities and functions.[25, 26] Here, we compare the binding and kinetic properties of CYP199A4 with selected cinnamic acid derivatives (Fig. 2). This will provide a better understanding of the structural features required for tight binding to this family of CYP enzymes. In addition the size and shape of substrates that can be efficiently oxidised by CYP199A4 will be determined. This will enable the engineering of CYP199A4 and analogous enzymes for improved substrate binding and turnover with non-natural substrates, enabling its use as a biocatalyst on a synthetically useful scale.

Results

Substrate binding studies on the CYP199A4 system

It has been reported that 4-hydroxycinnamic acid and 3-hydroxycinnamic acids are substrates for CYP199A2 from *R. palustris* CGA009 with the oxidation of both giving rise to 3,4-dihydroxycinnamic acid.[23, 24] The amino acid sequences and substrate binding profiles of the CYP199A2 and CYP199A4 enzymes are very similar.[4] The binding of both 3- and 4-hydroxycinnamic acids and the parent compound, cinnamic acid to CYP199A4, were studied. The addition of all of these substrates to CYP199A4 resulted in a spin state shift to the high-spin form of ≤5% (Table 1 and Fig. S1). 4-Hydroxycinnamic acid induced no observable spin state shift upon addition to CYP199A4. Due to the negligible shift in the spin state, an accurate determination of the dissociation constant was not possible for any of these substrates.

We have previously shown that a substituent *para* to the carboxylate moiety of benzoic acid was important for binding and efficient catalytic activity with both CYP199A2 and CYP199A4 [6, 20, 22] and 4-methoxybenzoic acid was the optimal substrate for both (≥95%, spin state shift and a K_d of 0.28 μM).[4, 5, 10] Accordingly, we tested the binding of 4-methoxycinnamic acid with CYP199A4. This substrate induced a spin state shift of 70% on binding to CYP199A4 and the dissociation constant was determined to be 3.6 μM (Table 1 and Fig. 3 and 4). 4-Methoxyphenylacetic acid binds poorly: spin state shift 5%; K_d 690 μM.[21] 3-(4-Methoxyphenyl)propionic acid induced a lower spin shift to the high-spin state (20%) and weaker binding (K_d, 31 μM) with CYP199A4 than both 4-methoxybenzoic and 4-methoxycinnamic acids. However the magnitude of this shift and the binding affinity is higher and than is observed with 4-methoxyphenylacetic acid and cinnamic acid (Table 1, Fig. 3 and 4).

Due to the favourable results obtained with 4-methoxycinnamic acid, we tested a range of substituted cinnamic acids with CYP199A4 to better determine the optimal size and shape of the molecules which complement the substrate binding pocket (Fig. 2). Additional substitutions at the

2- and 3-positions of 4-methoxycinnamic acid tended to reduce the affinity for binding to CYP199A4. 3-Hydroxy-4-methoxy-, 3,4-dimethoxy, 2,4-dimethoxy- and 3,4-methylenedioxy-cinnamic acids all induced significantly lower spin state shifts upon addition to CYP199A4 and resulted in weaker binding when compared to 4-methoxycinnamic acid (Table 1 and Fig. S1). 2,4-Dimethoxycinnamic acid bound the most tightly to CYP199A4 of any of the disubstituted substrates investigated despite possessing one of the lowest spin state shifts observed (Table 1). Substrates containing methoxy groups at the 3- or 5-positions with a hydrogen atom or a hydroxyl group at the 4-position such as 3,5-dimethoxycinnamic acid, 3-methoxy-4-hydroxycinnamic acid (ferulic acid) and 3,5-dimethoxy-4-hydroxycinnamic acid (sinapic acid) resulted in little observable shift in the spin state upon binding to CYP199A4 (Table 1 and Fig. S1). Two trimethoxy substituted cinnamic acids were tested with CYP199A4. Addition of neither 2,3,4-trimethoxycinnamic nor 3,4,5-trimethoxycinnamic acids induced a significant spin state shift (Table 1).

The effect of the size and type of the *para* substituent was also investigated with alkyl substituted cinnamic acids. Upon binding to CYP199A4, 4-methylcinnamic acid induced a similar spin state shift to 4-methoxycinnamic acid but the binding was 6-fold weaker (Table 1 and Fig. 3 and 4). The increased size of the alkyl *para*-substituent in 4-isopropylcinnamic acid resulted in the largest spin state shift observed, ≥ 95% high-spin and a comparable binding affinity to 4-methoxycinnamic acid (Table 1 and Fig. S2). The trend in the CYP199A4 substrate binding affinity for the alkyl and methoxy *para*-substituted cinnamic acids mirrors that of equivalent *para*-substituted benzoic acids (methoxy > isopropyl > methyl) despite the different trend in the spin state shifts.[5] On the whole, appropriately substituted cinnamic acids appear to be good target substrates for oxidation by CYP199A4.

Activity and product formation assays

The rate of NADH oxidation during the CYP199A4 catalysed oxidative demethylation of 4-methoxycinnamic acid was 560 nmol nmol-P450^{-1} min^{-1} (henceforth abbreviated to min^{-1}, Table 1 and Fig. 5). The product formation rate was inferior to that observed for the equivalent benzoic acid, predominantly due to diminished coupling efficiency of the NADH reducing equivalents to product formation (32% *vs.* 89-90%).[6] The oxidation of 4-methoxycinnamic acid by CYP199A4 resulted in the formation of a single product arising through demethylation of the methoxy group. This was identified by HPLC co-elution of a 4-hydroxycinnamic acid standard with the turnover (Fig. 6 and Scheme 1). We note that two peaks were observed in the HPLC and GC analysis of both the substrate and product controls of 4-methoxy- and 4-hydroxy-cinnamic acids. MS analysis indicated that these likely arise from the *E* and *Z* isomers (Fig. 6 and S4) with the starting, stable *E* isomer presumed to be present in excess in both instances. The NMR spectra of both substrate and product samples showed the presence of predominantly the *trans* isomer (>99%). Product samples isolated after whole-cell oxidation and HPLC purification had coupling constants for the alkene hydrogens consistent with a *trans* configuration (see supporting information). Overall isomerisation may be occurring before or during the chromatographic analysis of the turnovers.[27]

In accord with the reduced spin state shift and weaker binding of 3-(4′-methoxyphenyl)propionic acid to CYP199A4, the rate of NADH oxidation was lower than that for 4-methoxycinnamic acid (98 min^{-1}, Table 1 and Fig. 5). The coupling efficiency was higher than that observed with 4-methoxycinnamic acid and 4-methoxyphenylacetic acid, resulting in a product formation rate of 43 min^{-1} (Table 1). A single product was observed in the HPLC analysis of the turnover which co-eluted with 3-(4′-hydroxyphenyl)propionic acid (Scheme 1 and Fig. S3).

The addition of cinnamic acid and 3-hydroxycinnamic acid to CYP199A4 resulted in no appreciable increase in the rate of NADH oxidation above that observed in absence of substrate. 4-Hydroxycinnamic acid induced a very small increase in NADH consumption (Table 1). Chromatographic analysis of the CYP199A4 catalysed oxidation of these two substrates revealed

little product formation (Fig. S3) but 3-hydroxycinnamic acid did yield very low levels of 3,4-dihydroxycinnamic acid (Fig. S3). HPLC analysis of the cinnamic acid turnover showed trace amounts of 3- and 4-hydroxycinnamic acids, though most if not all of the 3-hydroxycinnamic acid could potentially be accounted for from impurities in the substrate (Fig. S3).

Minimal increases in the rate of NADH oxidation above that of the background NADH oxidase activity of the HaPux/HaPuR/CYP199A4 system, in the absence of substrate, were observed with 3-methoxy-4-hydroxy-, 3,5-dimethoxy- and 3,5-dimethoxy-4-hydroxy-cinnamic acids (Table 1). No product arising from substrate oxidation could be identified from CYP199A4 mediated oxidation of 3-methoxy-4-hydroxycinnamic acid or 3,5-dimethoxy-4-hydroxycinnamic acid, but very low levels of a single product were detected from 3,5-dimethoxycinnamic acid (Fig. S3). GC-MS analysis revealed a product with a mass consistent with 3-hydroxy-5-methoxycinnamic acid (expected mass: 338.6, observed mass: 338.3; Fig. S4 and S5), the product from a single oxidative demethylation.

Product formation from CYP199A4 catalysed oxidation of 3,4-(methylenedioxy)cinnamic was 30-and 50-fold slower than that observed with 4-methoxy- and 3,4-dimethoxy-cinnamic acids (Table 1). The lower amount of product formed from is mainly due to low coupling efficiency, 13% (Table 1). The sole product was identified as 3,4-dihydroxycinnamic acid arising from demethenylation (Fig. 6 and Scheme 1).

The oxidation of 3,4-dimethoxy-, 2,4-dimethoxy- and 3-hydroxy-4-methoxy-cinnamic acids by CYP199A4 resulted solely in the demethylation of the methoxy group at the 4-position and formation of a single product (Fig. S3 and S6). Despite the lower spin state shifts and weaker binding to the enzyme compared to 4-methoxycinnamic acid, the rate of NADH oxidation was high in all cases (Table 1). In fact, the levels of metabolite generated exceeded that of 4-methoxycinnamic acid due to higher coupling efficiencies (59-75%) and resulted in product formation rates in excess of 200 min^{-1} (Table 1).

The turnover of 3,4,5-trimethoxycinnamic acid by CYP199A4 did not result in any NADH oxidation activity above that of background NADH oxidase activity and no product formation was observed. 2,3,4-Trimethoxycinnamic was oxidised by CYP199A4 to generate a single product (Fig. 6 and Scheme 1). Although the NADH oxidation activity was low the coupling efficiency of this turnover was high resulting in a product formation rate of 24 min^{-1} (Table 1). GC-MS analysis indicated that the product arose from a single oxidative demethylation (substrate expected mass: 310.4, observed mass: 310.0; oxidative demethylation product; expected mass: 368.6, observed mass: 368.0; Fig. S4 and S5). The product was assigned as 4-hydroxy-2,3-dimethoxycinnamic acid based on the exclusive attack on the *para*-substituent reported so far for CYP199A4.

Substitution of the methoxy group with a methyl group resulted in a reduction in the NADH oxidation activity and product formation rate of the enzyme, 85 min^{-1} (Table 1). However, the coupling of the CYP199A4 turnover of 4-methylcinnamic acid was higher than that of 4-methoxycinnamic acid (Table 1). As a result, the relative reduction in activity reported is not as dramatic as that observed for the CYP199A4 catalysed turnover of 4-methylbenzoic acid compared to 4-methoxybenzoic acid.[6] A single product was observed in the CYP199A4 mediated oxidation of 4-methylcinnamic acid (Fig. 6). This was generated in higher yield using a whole-cell oxidation system (*vide infra*), purified via semi-prep HPLC and characterised via NMR as 4-hydroxymethylcinnamic acid (Scheme 1 and Fig. S6).

Despite the higher spin state shift of CYP199A4 with 4-isopropylcinnamic acid, the enzyme-catalysed oxidation of this substrate was slower than that of 4-methylcinnamic acid (Table 1). Four products were observed in the analysis of the *in vitro* turnovers by HPLC and GC-MS (Scheme 2 and Fig. S6). These products were also formed when using the whole-cell oxidation system, although in different proportions. The masses of the derivatised products indicated one resulted from dehydrogenation, two others from hydroxylation and another from further oxidation of the alkene or an alcohol product (Fig. S5 and S6 and Table S2).[6, 10, 28] The products were

generated using a whole-cell oxidation system and three were isolated and purified by HPLC. They were identified by NMR as 4-(1′-hydroxyisopropyl)cinnamic acid, 4-(2′-hydroxyisopropyl)cinnamic acid and 4-(1′,2′-epoxyisopropyl)cinnamic acid (Scheme 1 and Supporting information). The epoxide product presumably arises from further oxidation of 4-(prop-1′-en-2′-yl)cinnamic acid, the desaturation product of 4-isopropylcinnamic acid (Scheme 2).

Oxidation of cinnamic acid derivatives by whole-cell oxidation system

Previously, preparative product formation from 4-hydroxycinnamic and 3-hydroxycinnamic acids has been reported using whole-cell oxidation systems consisting of CYP199A2 and a combination of the electron transfer partners, palustrisredoxin, Pux, and putidaredoxin reductase.[23] Small scale (250 μL), high cell density (50 g cell wet weight L^{-1}) cultures were reported to generate 1 mM 3,4-dihydroxycinnamic acid from 4-hydroxycinnamic acid in 3 hours with the wild-type CYP199A2 enzyme. Larger scale turnovers (50 mL, 50 g cell wet weight L^{-1}) with the same system and substrate were reported to generate 3 mM product in 24 hours. In order to investigate the whole-cell oxidation of cinnamic acid substrates, we utilised a system developed in our laboratory which co-expresses CYP199A4, HaPux and HaPuR in *E. coli*.[4] Using larger volume (35 mL), low cell density cultures (3 g cell wet weight L^{-1}) and adding substrate (concentration 2 mM) with an additional equal aliquot being added at 4 hours, we found that all the substrates yielded the same products found in the *in vitro* turnovers (Scheme 1, Fig. 7 and Fig. S7).

In the CYP199A4 whole-cell turnovers of 4-methylcinnamic acid and 4-methoxycinnamic acid, all of the substrate (>95%) was converted to product at both 4 and 24 hours (Fig. 7 and S7). 3-Hydroxy-4-methoxycinnamic and 3-(4′-methoxyphenyl)propionic acids were the only other substrates turnovers where all of the material added (2 mM) was consumed within 4 hours (Fig. 7 and S7). However, only in the turnover of 3-hydroxy-4-methoxycinnamic acid was all the additional substrate consumed after 24 hours. Significant levels of the substrate (~20%) remained in the 3-(4-methoxyphenyl)propionic acid turnover after 24 hours (Fig. S7). In the turnover of 2,4-

dimethoxycinnamic acid, most of the substrate (~75%) was depleted and the product formation was high after 4 hours but a higher ratio of substrate to product was observed after 24 hours (Fig. S7). The turnover of 3,4-dimethoxycinnamic acid resulted in conversion of approximately half the substrate in the first 4 hours (~960 μM 3-methoxy-4-hydroxycinnamic acid), and a similar proportion of starting material and product was observed at 24 hours (Fig. S7). Oxidation of 4-methoxyphenylacetic acid by CYP199A4 generated the next highest level of product with 320 and 510 μM of 4-hydroxyphenylacetic acid being obtained after 4 and 24 hours, respectively (Fig. S7).

The levels of product in the turnovers in which the substrate lacks an alkyl or methoxy *para*-substituent, were very low (≤ 120 μM after 24 hours). No monooxygenase product was observed at either 4 or 24 hours with cinnamic, 3-hydroxycinnamic and 4-hydroxy-3,5-dimethoxycinnamic acids (Fig. S7). There was no demethylated metabolite in the turnover of 3,5-dimethoxycinnamic acid after 4 hours but low levels of product (65 μM) were observed after 24 hours (Fig. S7). Low levels of 3,4-dihydroxycinnamic acid were found in the whole-cell oxidation of 4-hydroxycinnamic acid at both 4 and 24 hours (~30 and 120 μM, respectively, Fig. S7).

Discussion

The tight binding of 4-methoxycinnamic acid to CYP199A4 was somewhat surprising as while 4-methoxybenzoic acid binds well to this enzyme small changes at the carboxylate group have been shown to result in a large reduction in binding affinity with even 4-methoxyphenylacetic acid being a poor substrate. The planar nature of 4-methoxycinnamic acid seems to be important for substrate binding. As with 4-methoxybenzoic, 2-naphthoic and indole-6-carboxylic acids the planarity of the substrate may maintain the interactions of the carboxylate group with the arginine and serine residues as well as the hydrophobic contacts (Fig. 1).[5, 6, 16, 18] Overall, the results with the cinnamic acids reinforce the importance of the presence and the identity of the functional group at the *para*-position. The trends in the activity and product selectivity of CYP199A4 broadly agree with those obtained with similarly substituted methoxybenzoic acid substrates with low amounts of product being formed in the absence of a 4-methoxy substituent. For example, methoxy and alkyl groups are preferred over a hydroxyl moiety or no substitution at all.

The turnover of 4-isopropylcinnamic acid generated products arising from hydroxylation at both the α- and β-carbons. Metabolites arising from desaturation pathways were also observed, in agreement with the products reported from the oxidation of 4-isopropylbenzoic acid by CYP199A4. Where the cinnamic acid substrate lacked a substituent or contained a hydroxy group at the *para*-position the turnovers generated very low levels of product highlighting the selectivity of the enzyme for this position. It may be the case that substrates containing a *para*-hydroxy substitution are close enough coordinate to the heme-iron which may interfere with the observation of a type I spin state shift and the commencement of the P450 catalytic cycle. In the turnover of 3,5-dimethoxybenzoic acid by CYP199A4, a small amount of 3-hydroxy-5-methoxybenzoic acid was observed. It has been hypothesised that this arises from a substrate methoxy group being oriented toward the heme-iron and therefore close enough to be oxidised.[20] If only one substituent is present

at the 3-position, it is likely to point away from the heme in a similar orientation to that observed in the crystal structure of 3,4-dimethoxybenzoic acid-bound CYP199A4 (PDB code; 4EGN).[5]

The impact of additional substitutions in the cinnamic acid structure on substrate binding to and the activity of CYP199A4 is generally representative of what was previously observed with substituted benzoic acids.[4, 5, 20, 22] However, the punitive effect of the additional substitutions at the 3-position of the cinnamic acids was greater than the equivalent modification of benzoic acids and this may be related to the larger size of the cinnamic acids. The larger trimethoxycinnamic acids seem to be too bulky for the substrate binding pocket of CYP199A4 or their shape many force the substrate to bind in such a fashion that the water bound to the heme iron is not displaced.

Overall, the trend in the levels of product formed in the *in vitro* CYP199A4 turnovers were mirrored in the whole-cell assays. Complete conversion of suitable substrates at concentrations up to 4 mM could be achieved using low cell density whole-cell oxidation in shake flasks. There was higher than expected levels of product observed in the whole-cell turnovers of 4-hydroxycinnamic acid when compared to the *in vitro* turnover studies. Recent work has shown that the *E. coli* hydroxylase complex, HpaBC, is capable of oxidising this substrate to 3,4-dihydroxycinnamic acid which may account for this.[29]

The selectivity of CYP199A4 oxidation of the cinnamic acids for the *para*-substituent is interesting as chemical oxidants might be expected to target the more reactive alkene group. Indeed chloroperoxidase oxidises cinnamic acid and its derivatives at the alkene group.[30] Cinnamic acids and their derivatives are phenylpropanoids which are naturally found in plants. They are constituent components of the biosynthetic pathways of many natural products encompassing coumarins and flavonoids.[25, 26] They have a range of medicinal properties including anticancer, antimicrobial and antioxidant activities.[25] Therefore new biocatalytic methods to facilitate the synthesis of cinnamic derivatives could enhance their use and effectiveness. For example Ozagrel, an antiplatelet agent inhibiting thromboxane A2, is related to 4-methylcinnamic acid.[31] CYP199A4 could have a role as

a biocatalyst for the modification of substituted cinnamic acid derivatives and the oxidation of these substrates could be improved using protein engineering techniques to optimise the activity. For example the F185L mutant of CYP199A2 is reported to give increased levels of product with 4-hydroxycinnamic acid and is capable of oxidising cinnamic acid to 3,4-dihydroxycinnamic acid.[23]

Experimental

General reagents, organic substrates and *N*,*O*-bis(trimethylsilyl)trifluoroacetamide with trimethylsilyl chloride (BSTFA + TMSCl, 99:1) were from Sigma-Aldrich, TCI, Fluorochem or VWR. Buffer components, NADH and isopropyl-β-D-thiogalactopyranoside (IPTG) were from Astral Scientific, Australia. UV/Vis spectra and spectroscopic activity assays were recorded at 30 ± 0.5 °C on an Agilent CARY-60 or Varian CARY-5000 spectrophotometer. Analytical High Performance Liquid Chromatography (HPLC) was performed on an Agilent 1260 Infinity Pump equipped with an autoinjector connected using an Agilent Eclipse Plus C18 column (250 mm x 4.6 mm, 5 μm) or a Jupiter C18 300A column (250 x 4.6 mm, 5 μm). The products were separated using a gradient between 20-95%, 20-50% or 0-50% acetonitrile in water (0.1% trifluoroacetic acid (TFA) at a flow rate of 1 mL min^{-1} over 30 minutes. Gas Chromatography-Mass Spectrometry (GC-MS) data were collected on a Shimadzu GC-17A using a QP5050A GC-MS detector and a DB-5 MS fused silica column (30 m x 0.25 mm, 0.25 μm). The injector and interface were maintained at a constant temperature of 250 °C and 280 °C. The oven temperature was held at 100 ºC for 1 min and then increased at 15 ºC min^{-1} up to 220 ºC. The HPLC retention times for the substrates and products and the GC retention times of their trimethylsilyl (TMS) derivatives are given in the supporting information (Table S1 and S2).

The expression and purification of CYP199A4, HaPux and HaPuR have been described elsewhere.[4-6] The CYP199A4 protein concentration was calculated using ε_{419} = 119 mM^{-1} cm^{-1}.[5]

Substrate binding: spin state determination and binding titrations

The high-spin heme content was determined using an enzyme concentration of 1.8 – 3.7 μM in 50 mM Tris, pH 7.4, with the addition of small aliquots of substrate (from a 100 mM stock) until the spectrum did not change. The percentage shift was estimated (to approximately ±5%) by comparison with a set of spectra generated from the sum of the appropriate percentages of the

spectra of the substrate-free (>95% low-spin, Soret maximum at 418 nm) and camphor-bound (>95% high-spin, Soret maximum at 392 nm) forms of WT CYP101A1.

For dissociation constant determination CYP199A4 was diluted to 1.3 – 5.5 μM using 50 mM Tris, pH 7.4, in 2.5 mL and 0.5 – 2 μL aliquots of the substrate were added using a Hamilton syringe from 1, 10 or 100 mM stock solutions in ethanol or DMSO (Fig. 3 and S2). The maximum Soret band peak-to-trough difference (ΔA) in absorbance was recorded between 700 nm and 250 nm. Further aliquots of substrate were added until the peak-to-trough difference of the Soret band did not change. The dissociation constants, K_d, were obtained by fitting ΔA against total substrate concentration [S] to a hyperbolic function:

$$\Delta A = \frac{\Delta A_{\max} \times [S]}{K_d + [S]}$$

where $\Delta A_{\max}$ is the maximum absorbance difference. 4-Methoxy- and 4-isopropyl-cinnamic acids exhibited tight binding, with K_d < 5 μM and their data were fitted to the tight binding quadratic equation:[32]

$$\frac{\Delta A}{\Delta A_{\max}} = \frac{([E]+[S]+K_d) - \sqrt{\{([E]+[S]+K_d)^2 - 4[E][S]\}}}{2[E]}$$

where $\Delta A_{\max}$ is the maximum absorbance difference and [E] is the enzyme concentration.

Activity assays

In vitro NADH turnover rate assays were performed with mixtures (1.2 mL) containing 50 mM Tris, pH 7.4, 0.5 μM CYP199A4, 5 μM HaPux and 0.5 μM HaPuR. The buffer solution was oxygenated before use and the mixtures equilibrated at 30 °C for 2 min. Substrates were added as a 100 mM stock solution in ethanol or DMSO to a final concentration of 1 mM. NADH was added to *ca*. 320 μM (final A_{340} = 2.00) and the absorbance at 340 nm was monitored. The rate of NADH oxidation was calculated using ε_{340} = 6.22 $mM^{-1}cm^{-1}$. For 3,4-(methylenedioxy)cinnamic acid which absorbed at 340 nm, the reaction was monitored at 370 nm using ε_{370} = 2.59 $mM^{-1}cm^{-1}$.

In order to isolate and identify products for which no standards were available, and to assess the activity of the CYP199A4 system in *E. coli*, we utilised a whole-cell oxidation system comprising of the plasmids pETDuetHaPux/HaPuR and pRSFDuetHaPux/CYP199A4, the construction and use of which has been described previously.[4] The cell pellet from a 100 ml growth (~ 6 g cell wet weight L^{-1}) was resuspended in double the volume of *E. coli* minimal media (EMM).[33] For small scale turnovers this cell suspension was split into 30 mL aliquots in 250 mL Erlenmeyer flasks. The substrates were added to the resuspended cells to a concentration of 2 mM and the reactions were then shaken at 30 °C and 200 rpm. Samples for HPLC analysis were taken after 4 hours at which point a further 2 mM aliquot of substrate was added and a second sample was taken after 24 hours. The supernatant was separated from the cells by centrifugation before analysis.

For large scale growths 200 mL of the cell suspension was added to a 2 L baffled flask and 2 mM aliquots of substrate were added at 1, 3 and 6 hours. After 20 hours, the supernatant (200 mL) was acidified, extracted in ethyl acetate (3 x 100 mL), washed with brine (100 mL) and dried with $MgSO_4$. The organic extracts were pooled and the solvent was removed by vacuum distillation and then under a stream of nitrogen. The products were purified using an Agilent 1100 HPLC equipped with Supelcosil LC-18 semi-prep column (5 μm particle size, 25 cm × 10 mm) and a fraction collector. A gradient, 20 - 50% of acetonitrile (with trifluoroacetic acid, 0.1%) in water (TFA, 0.1%) was used with UV detection at 240, 254 and 280 nm. Those fractions containing a single product (≥95%) were combined for characterisation. The solvent was removed under reduced pressure and the purified product was dissolved in deuterated DMSO. NMR spectra were acquired on an Agilent DD2 spectrometer operating at 500 MHz for ^{1}H and 126 MHz for ^{13}C. A combination of ^{1}H and ^{13}C experiments were used to determine the structures of the products (see supporting information).

Analysis of metabolites

The majority of the products from the enzyme turnovers were identified via co-elution of authentic standards (Scheme 1) using HPLC analysis or via GC-MS of the trimethylsilyl chloride (TMSCl) derivatised forms.[4] After the NADH had been consumed in substrate oxidation incubations, 150 μL of the reaction mixture was mixed with 2 μL of an internal standard solution (20 mM 9-hydroxyfluorene in ethanol). This was mixed with 50 μL of acetonitrile before analysis by HPLC. The HPLC retention times for the substrates and products are given in the supporting information (Table S1).

For gas chromatography analysis, 990 μL of the reaction mixture was mixed with 10 μL of an internal standard solution (20 mM 9-hydroxyfluorene) and 3 μL of 3 M HCl. The mixture was extracted three times with 400 μL of ethyl acetate and the organic extracts were combined and dried over $MgSO_4$. Solvent was evaporated under a stream of nitrogen and the sample dissolved in 150 μL acetonitrile. Excess (35 μL) BSTFA + TMCS (99:1) was added and the mixture left for at least 120 min to produce the trimethylsilyl ester of the carboxylic acid group and trimethylsilyl ether of the alcohol, if formed. These reaction mixtures were used directly for GC-MS analysis. The retention times for the trimethylsilyl (TMS) derivatives are given in the supporting information (Table S2). Calibrations of fixed amounts of these standards were used to quantitate the level of product in the turnovers by HPLC or GC-MS.

Conclusion

The CYP199A4 product formation activity varied considerably across the range of cinnamic acid derivatives. 4-Methoxycinnamic acid was a better substrate than 3-(4-methoxyphenyl)propionic acid which was superior to 4-methoxyphenylacetic acid. Suitably *para*-substituted cinnamic acids were excellent substrates and their turnover was selective for oxidation at the *para*-position. These substrates were also found to bind with high affinity. Whole-cell oxidation systems were used to generate significant levels of product for 3-(4-methoxyphenyl)propionic acid and similarly substituted cinnamic acids. The regioselectivity was high with the substituent at the 4-position being the sole site of oxidation when a suitable functional group was present. This selectivity could be useful in synthetic applications, such as selective hydroxylation and/or deprotection reactions. As a result the biocatalytic scope of CYP199A4 has been expanded and this knowledge can be used to improve substrate binding and turnover of this enzyme for non-natural substrates.

Acknowledgment

This work was supported by ARC grant DP140103229 (to JJDV and SGB). SGB acknowledges the ARC for a Future Fellowship (FT140100355). The authors also thank the University of Adelaide for the award of a M. Phil Scholarship to RRC.

References

1. P. R. Ortiz de Montellano, ed., *Cytochrome P450: Structure, Mechanism, and Biochemistry* Kluwer Academic/Plenum Press, New York, 2005.
2. P. R. Ortiz de Montellano, *Chem. Rev.*, 2010, **110**, 932-948.
3. A. Sigel, H. Sigel and R. Sigel, eds., *The Ubiquitous Roles of Cytochrome P450 Proteins*, John Wiley & Sons, Weinheim, 2007.
4. S. G. Bell, A. B. Tan, E. O. Johnson and L. L. Wong, *Mol. Biosyst.*, 2010, **6**, 206-214.
5. S. G. Bell, W. Yang, A. B. Tan, R. Zhou, E. O. Johnson, A. Zhang, W. Zhou, Z. Rao and L. L. Wong, *Dalton Trans.*, 2012, **41**, 8703-8714.
6. S. G. Bell, R. Zhou, W. Yang, A. B. Tan, A. S. Gentleman, L. L. Wong and W. Zhou, *Chemistry*, 2012, **18**, 16677-16688.
7. S. G. Bell, J. H. McMillan, J. A. Yorke, E. Kavanagh, E. O. Johnson and L. L. Wong, *Chem. Commun.*, 2012, **48**, 11692-11694.
8. F. Xu, S. G. Bell, Y. Peng, E. O. Johnson, M. Bartlam, Z. Rao and L. L. Wong, *Proteins*, 2009, **77**, 867-880.
9. J. Rittle and M. T. Green, *Science*, 2010, **330**, 933-937.
10. S. G. Bell, F. Xu, I. Forward, M. Bartlam, Z. Rao and L.-L. Wong, *J. Mol. Biol.*, 2008, **383**, 561-574.
11. F. P. Guengerich, *Chem. Res. Toxicol.*, 2001, **14**, 611-650.
12. S. G. Bell, N. Hoskins, C. J. C. Whitehouse and L. L. Wong, in *Metal Ions in Life Sciences*, eds. A. Sigel, H. Sigel and R. Sigel, John Wiley & Sons, 1st edn., 2007, vol. 3, Chapter 14, pages 437-476, p. 652.
13. V. B. Urlacher and M. Girhard, *Trends Biotechnol.*, 2012, **30**, 26-36.
14. C. J. Whitehouse, S. G. Bell and L. L. Wong, *Chem. Soc. Rev.*, 2012, **41**, 1218-1260.
15. M. T. Lundemo and J. M. Woodley, *Appl. Microbiol. Biotechnol.*, 2015, **99**, 2465-2483.
16. T. Furuya and K. Kino, *Biosci. Biotechnol. Biochem.*, 2009, **73**, 2796-2799.
17. T. Furuya and K. Kino, *ChemSusChem*, 2009, **2**, 645-649.
18. T. Furuya and K. Kino, *Appl. Microbiol. Biotechnol.*, 2010, **85**, 1861-1868.
19. T. Furuya, Y. Shitashima and K. Kino, *J. Biosci. Bioeng.*, 2015, **119**, 47-51.
20. T. Coleman, R. R. Chao, J. B. Bruning, J. De Voss and S. G. Bell, *RSC Adv.*, 2015, **5**, 52007 - 52018.
21. T. Coleman, R. R. Chao, J. De Voss and S. G. Bell, *Biochim. Biophys. Acta Proteins and Proteomics*, 2016, **1864**, 667-675.
22. S. G. Bell, N. Hoskins, F. Xu, D. Caprotti, Z. Rao and L. L. Wong, *Biochem. Biophys. Res. Commun.*, 2006, **342**, 191-196.
23. T. Furuya, Y. Arai and K. Kino, *Appl. Environ. Microbiol.*, 2012, **78**, 6087-6094.
24. J. L. Rodrigues, R. G. Araújo, K. L. J. Prather, L. D. Kluskens and L. R. Rodrigues, *Enzyme Microb. Tech.*, 2015, **71**, 36-44.
25. J. D. Guzman, *Molecules*, 2014, **19**, 19292-19349.
26. O. Yu and J. M. Jez, *Plant J.*, 2008, **54**, 750-762.
27. Y. S. Rao and R. Filler, *Chem. Commun.*, 1976, 471-472.
28. C. J. Whitehouse, S. G. Bell and L. L. Wong, *Chemistry*, 2008, **14**, 10905-10908.
29. J. A. Jones, S. M. Collins, D. M. Lachance, V. R. Vernacchio and M. A. Koffas, *Biotechnol. Prog.*, 2015.
30. H. Yamada, N. Itoh and Y. Izumi, *J. Biol. Chem.*, 1985, **260**, 11962-11969.
31. J. M. Dogne, X. de Leval, P. Benoit, S. Rolin, B. Pirotte and B. Masereel, *Expert Opin. Investig. Drugs*, 2002, **11**, 275-281.
32. J. W. Williams and J. F. Morrison, *Methods Enzymol.*, 1979, **63**, 437-467.
33. S. G. Bell, C. F. Harford-Cross and L. L. Wong, *Protein Eng.*, 2001, **14**, 797-802.

Tables

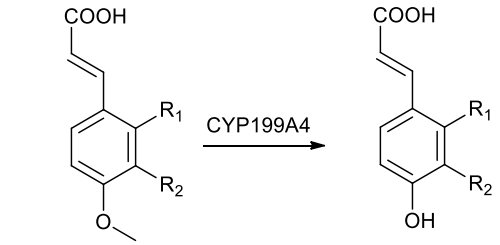
Table 1 Substrate binding parameters and catalytic turnover activity data for CYP199A4 with various substrates. The data are given as mean ± S.D. with $n \geq 3$. The reaction mixtures (in 50 mM Tris, pH 7.4) contained 0.5 μM P450, 5 μM HaPux and 0.5 μM HaPuR. Rates are given as nmol nmol-CYP^{-1} min^{-1}. The average background NADH consumption rate was 9.0 nmol nmol-CYP^{-1} min^{-1}. The data for 4-methoxyphenylacetic acid were reported previously and are included for comparison.[21]

Substrate	**% HS**	**K_d/μM**	**NADH**	**PFR[a]**	**C[b] %**
4-methoxycinnamic acid	70%	3.6 ± 0.1	560 ± 10	180 ± 9	32 ± 1
cinnamic acid	<5%	-[c]	24 ± 1	-[e]	-[e]
3-(4′-methoxyphenyl)propionic acid	20%	31 ± 1	98 ± 5	43 ± 3	43 ± 2
4-methylcinnamic acid	70%	21 ± 0.4	202 ± 7	85 ± 1	42 ± 1
3-hydroxycinnamic acid	<5%	-[c]	9 ± 0.1	-[e]	-[e]
4-methoxyphenylacetic acid[21]	5%	690 ± 70	21 ± 1.5	5.0 ± 0.6	23 ± 2
4-hydroxycinnamic acid	0%	-[c]	17 ± 0.5	0.6 ± 0.04	3.4 ± 0.2
3,4-(methylenedioxy)cinnamic acid	30%	120 ± 2	51 ± 0.4 [f]	6.6 ± 0.2	13 ± 0.5
4-isopropylcinnamic acid	>95%	3.4 ± 0.1	190 ± 13	131 ± 10	69 ± 3
3,5-dimethoxy-4-hydroxycinnamic acid	<5%	-[c]	50 ± 2	-[g]	-[g]
3,5-dimethoxycinnamic acid	<5%	-[c]	22 ± 1	3.3 ± 0.2	0.3 ± 0.02
3-methoxy-4-hydroxycinnamic acid	<5%	-[c]	-[d]	-[g]	-[g]
3-hydroxy-4-methoxycinnamic acid	25%	224 ± 13	408 ± 25	239 ± 5	59 ± 3
3,4-dimethoxycinnamic acid	20%	840 ± 30	400 ± 20	302 ± 17	75 ± 4
2,4-dimethoxycinnamic acid	10%	86 ± 4.5	441 ± 30	282 ± 25	64 ± 2
2,3,4-trimethoxycinnamic acid	<5%	-[c]	39 ± 2	24 ± 1	61 ± 3
3,4,5-trimethoxycinnamic acid	<5%	-[c]	-[d]	-[g]	-[g]

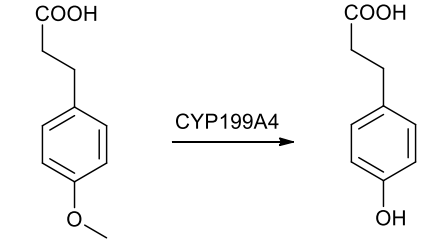
[a] PFR: product formation rate. [b] C: the coupling efficiency; the percentage of NADH consumed in the reaction that led to the formation of products. [c] Not determined due to absence of spin state shift. [d] No significant NADH turnover activity was observed. [e] Minimal levels of product formation were observed (Fig. S3) [f] Determined at A_{370} due to interfering absorption of the substrate. [g] no significant product observed.

Figures and Schemes

Scheme 1 The products formed from CYP199A4 enzyme turnovers with different cinnamic acid derivatives (abbreviations; CA, cinnamic acid; MCA, methoxycinnamic acid).

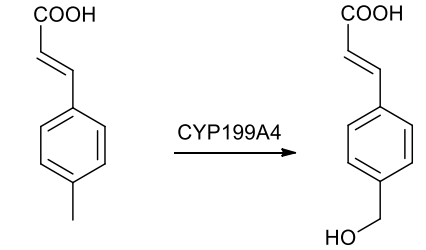


Substrate	R_1	R_2
4-methoxyCA	H	H
3,4-dimethoxyCA	H	OMe
2,4-dimethoxyCA	OMe	H
3-hydroxy-4-MCA	H	OH
2,3,4-trimethoxyCA	OMe	OMe



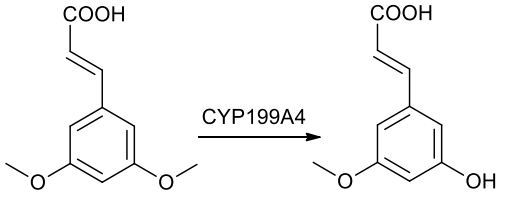
3-(4'-methoxyphenyl)-propionic acid

3-(4'-hydroxyphenyl)-propionic acid



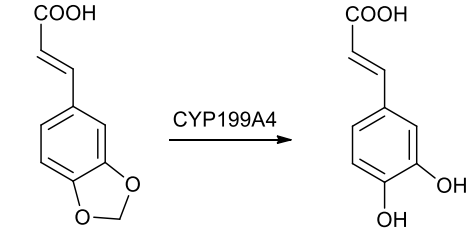
4-methyl-cinnamic acid

4-hydroxymethyl-cinnamic acid



3,5-dimethoxy-cinnamic acid

3-hydroxy-5-methoxy-cinnamic acid



3,4-(methylene-dioxy)cinnamic acid

3,4-dihydroxy-cinnamic acid

Scheme 2 The products formed from CYP199A4 enzyme turnovers with 4-isopropylcinnamic acid The products were identified as; A, 4-(1′-hydroxyisopropyl)cinnamic acid, B, 4-(2′-hydroxyisopropyl)cinnamic acid, C, 4-(prop-1′-en-2′-yl)cinnamic and D, 4-(1′,2′-epoxyisopropyl)cinnamic acid.

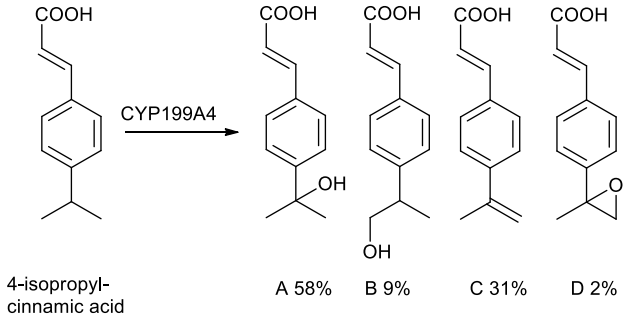


Figure 1 The active site of 4-methoxybenzoic acid-bound CYP199A4 (PDB code 4DO1). The amino acid residues which confer the substrate-specificity of the enzyme have been shown. 4-Methoxybenzoic acid is shown in yellow, the heme in grey and the amino acids in green.

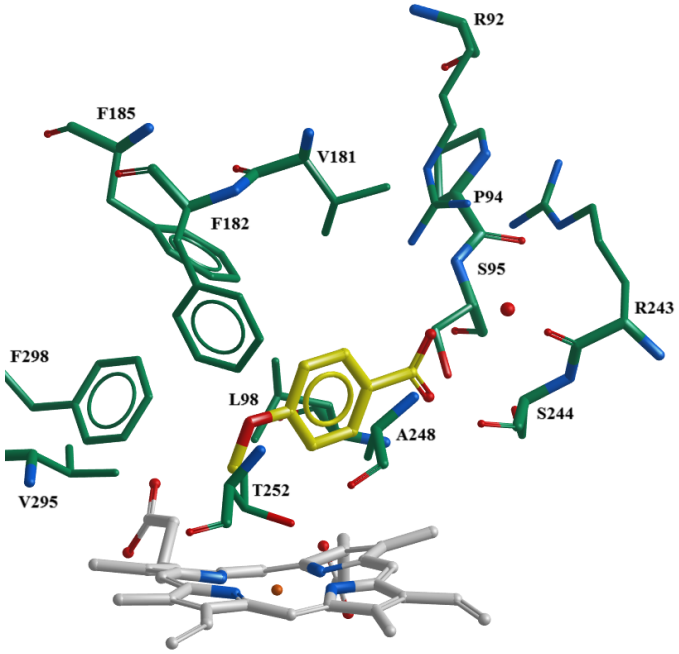


Figure 2 The substrates tested for binding and activity with CYP199A4.

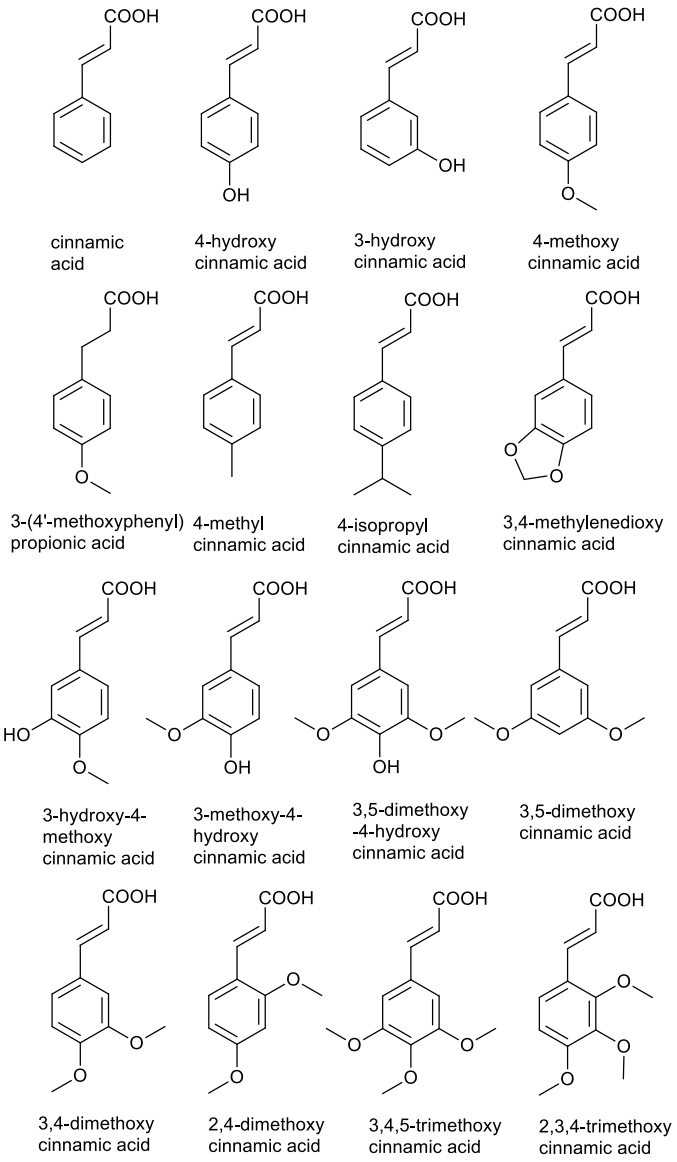


Figure 3 Spin state shift of CYP199A4 with the substrates (a) 3-(4′-methoxyphenyl)propionic acid, (b) 4-methoxycinnamic acid, (c) 4-methylcinnamic acid and (d) 4-isopropylcinnamic acid. Red; substrate-bound CYP199A4, black; substrate-free CYP199A4.

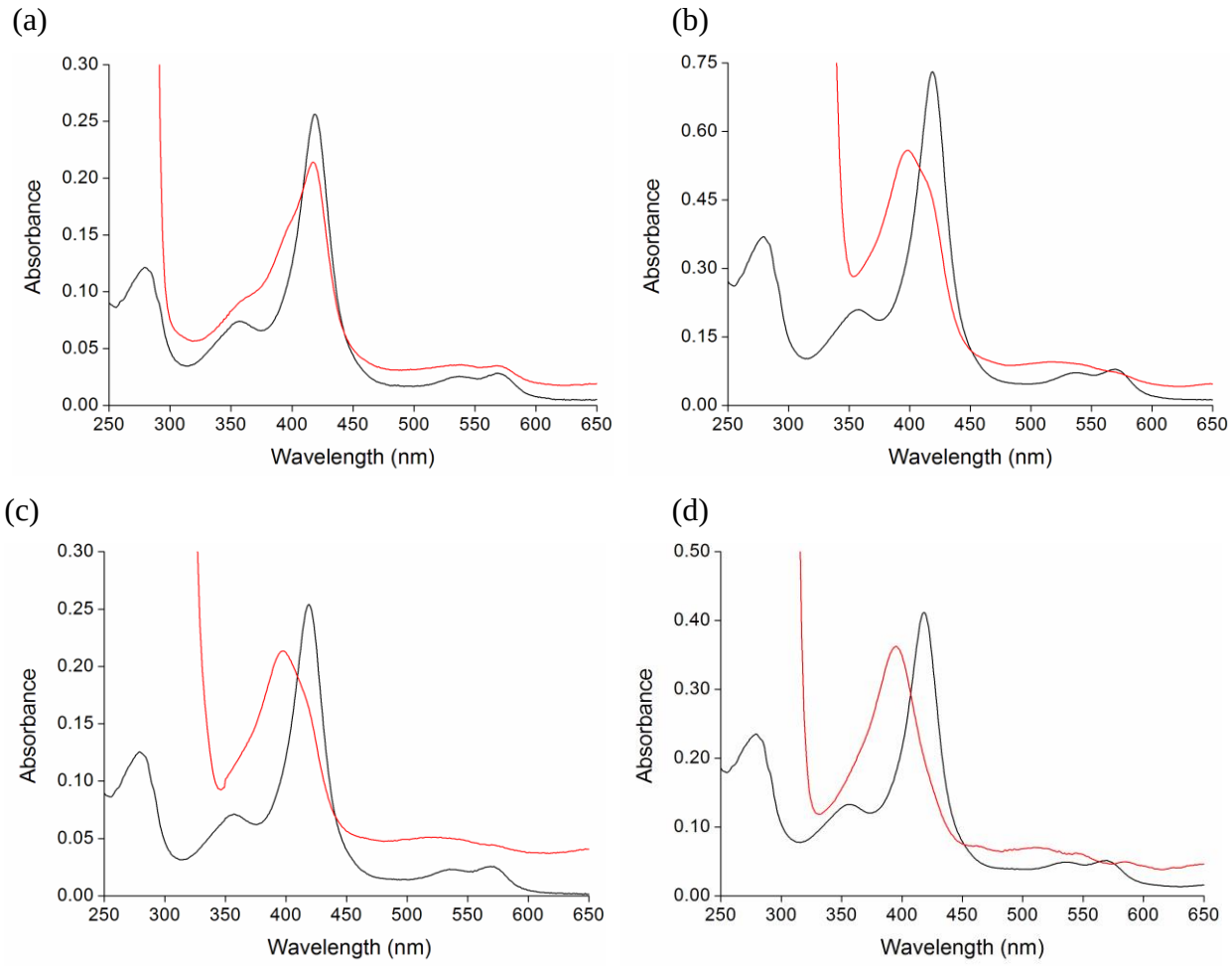


Figure 4 Dissociation constant analyses of selected substrates with CYP199A4. The protein concentration, the absorbance of the peak and trough used to calculate ΔA and the K_d are also provided. (a) 3-(4′-methoxyphenyl)propionic acid (CYP199A4 concentration, 2.1 μM, A_{390}-A_{421} and K_d 31 μM) (b) 4-methoxycinnamic acid (CYP199A4 concentration, 1.3 μM, A_{390}-A_{421} and K_d 3.6 μM) and (c) 4-methylcinnamic acid (CYP199A4 concentration, 1.5 μM, A_{390}-A_{421} and K_d 21 μM).

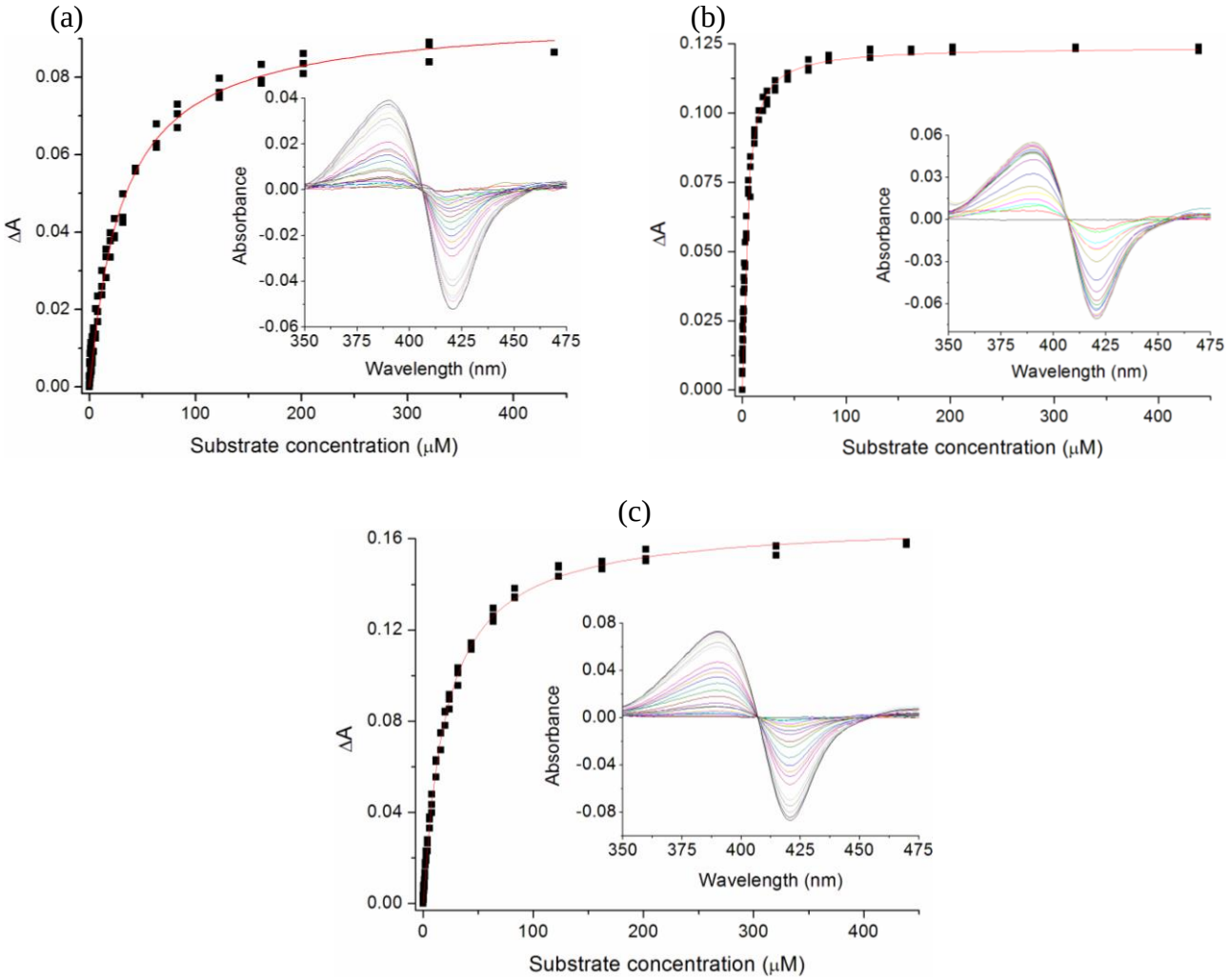


Figure 5 NADH oxidation assays of 4-methoxycinnamic acid (red), 4-methylcinnamic acid (black), 3-(4′-methoxyphenyl)propionic acid (green), 4-methoxyphenylacetic acid (blue), cinnamic acid (cyan), 4-hydroxycinnamic acid (purple) and 4-methoxybenzoic acid (magenta). See experimental section for details and Table 1.

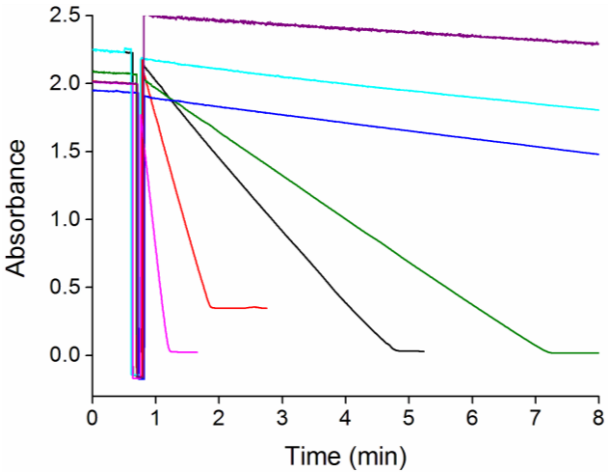
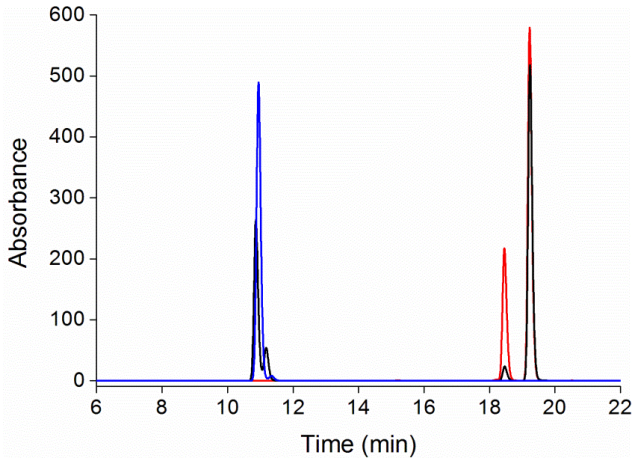
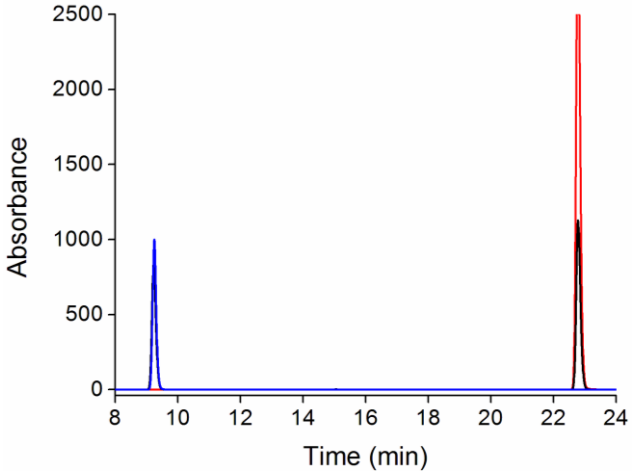


Figure 6 HPLC analysis of the *in vitro* turnovers of cinnamic acid derivatives. (a) HPLC analysis of the 4-methoxycinnamic acid turnover by CYP199A4 (black), 4-methoxycinnamic acid control (red) and 4-hydroxycinnamic acid control (blue). 4-Hydroxycinnamic acid (RT 11.0 min) and 4-methoxycinnamic acid 18.5 and 19.3 min (*cis* and *trans* isomers, respectively). (b) HPLC trace of the 4-methylcinnamic acid turnover by CYP199A4 (black), 4-methylcinnamic acid control (red) and 4-hydroxymethylcinnamic acid control (blue). 4-Hydoxymethylcinnamic acid (RT 9.3 min) and 4-methylcinnamic acid, RT 22.8 min. (c) HPLC trace of the 2,3,4-trimethoxycinnamic acid turnover by CYP199A4 (black), 2,3,4-trimethoxycinnamic acid control (red). The product assigned as 4-hydroxy-2,3-dimethoxycinnamic acid, RT 11.0 min, 2,3,4-trimethoxycinnamic acid, RT 17.2 min and internal standard, RT 22.1 min.

(a)



(b)



(c)

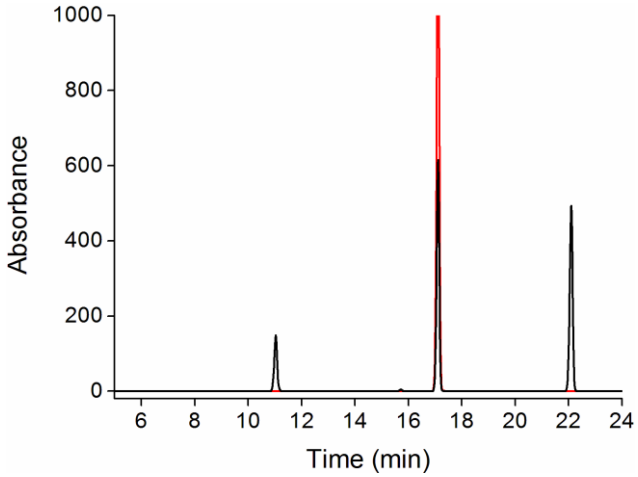
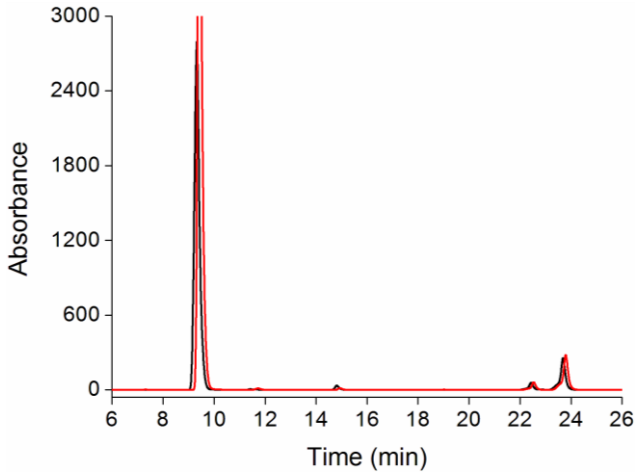
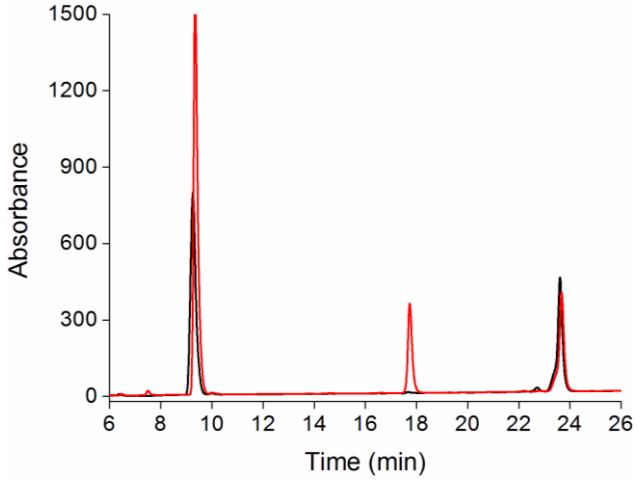

Absorbance
1000
800
600
400
200
0
6
8
10
12
14
16
18
20
22
24
Time (min)

Figure 7 HPLC analysis of the whole-cell oxidation of cinnamic acid derivatives; turnover at 4 hours (black), turnover overnight (red). 200 µM 9-hydroxyfluorene standard RT 23.8 min. (a) 4-Methylcinnamic acid – substrate RT 22.8 min, product RT 9.3 min. (b) 3-(4′-methoxyphenyl)propionic acid – substrate RT, 18.0 min, product, RT 9.6 min. (c) 4-methoxycinnamic acid – substrate RT 17.9 (*cis*) and 18.1 (*trans*) mins, product RT 11.0 min. The additional peak at 7.5 min is consistent with further oxidation of the 4-hydroxycinnamic acid product to 3,4-dihydroxycinnamic acid (Table S1).

(a)



(b)



(c)